



The Abdus Salam
**International Centre
for Theoretical Physics**
50th Anniversary 1964–2014



Indico built-in search plugin: **Repozer**

Giorgio Pieretti
pieretti@ictp.it

What is Repozer

- Indico plug-in that wraps repoze.catalog
- “repoze.catalog is a Python indexing and searching framework”





The Abdus Salam
**International Centre
for Theoretical Physics**
50th Anniversary 1964–2014



A quick look @ **Repoze.catalog**

How does it work – the Class

Example: How to index an Object



```
class City(Persistent):  
    def __init__(self, pop):  
        self.name = name  
        self.population = pop
```

How does it work – the catalog

```
_initialized = False

def initialize_catalog():
    '''
    factory = FileStorageCatalogFactory('../data/mdcatalog.db', 'mycatalog')

    NB: Use of global variable
    '''
    global _initialized
    if not _initialized:
        # create a catalog
        manager = ConnectionManager()
        catalog = factory(manager)
        # set up indexes
        catalog['names'] = CatalogTextIndex('name')
        catalog['populations'] = CatalogFieldIndex('population')
        # commit the indexes
        manager.commit()
        manager.close()
        _initialized = True
```

How does it work – the indexing

```
if __name__ == '__main__':  
    initialize_catalog()  
    manager = ConnectionManager()  
    catalog = factory(manager)  
  
    #For ease of demonstration set up a local dict  
    #containing a number of `City` instances keyed  
    #by a unique integer  
    cities = {  
        1:City('Windhoek', 322500),  
        2:City('Pretoria', 525387),  
        3:City('Nairobi', 3138295),  
        4:City('Maputo', 1244227),  
        5:City('Jakarta', 10187595),  
        6:City('Canberra', 358222),  
        7:City('Wellington', 393400),  
        8:City('Santiago', 5428590),  
        9:City('Buenos Aires', 2891082),  
    }  
  
    #Iterate over our local dict and for each  
    #element generate the catalog entry for  
    #repose.catalog and add the corresponding  
    #instance to the ZODB database we are  
    #cataloguing  
    for docid, doc in cities.items():  
        catalog.index_doc(docid, doc)
```

How does it work – the query

```
-  
print "Looking for 'less than' value on the `population` index"  
print "Populations less than 1,000,000"  
numdocs, results = catalog.query(Lt('populations', 1000000))  
print "Raw Result: "  
print (numdocs, [ x for x in results ])
```

```
print "Looking for 'InRange' values on the `population` index"  
print "Populations between 1,000,000 and 4,000,000"  
numdocs, results = catalog.query(InRange('populations',  
                                         1000000, 4000000))  
print "Raw Result: "  
print (numdocs, [ x for x in results ])
```

```
1 from repoze.catalog.query import Eq  
2 results = catalog.query(Eq('index_name', 'value'))
```

Or you can allow `repoze.catalog` to construct a query object on your behalf by passing a *string* as `queryobject`.

```
1 from repoze.catalog.query import Eq  
2 catalog.query('index_name == "value"')
```

```
1 from repoze.catalog.query import And, Eq, Contains  
2 query = And(Eq('author', 'crossi'), Contains('body', 'biscuits'))
```

How does it work - the storage

- External
 - repoze_catalog.fs
- Internal
 - Data.fs

```
from repoze.catalog.catalog import FileStorageCatalogFactory
from repoze.catalog.catalog import ConnectionManager
```

```
factory = FileStorageCatalogFactory(DBpath, 'repoze_catalog')
manager = ConnectionManager()
catalog = factory(manager)
```

```
dbc = db.DBMgr.getInstance().getDBConnection()
root = dbc.root()
if 'repozecatalog' not in root:
    init_catalog()
    catalog = root['repozecatalog']
```



Have to deal with ConnectionManager
to avoid “Lock error”



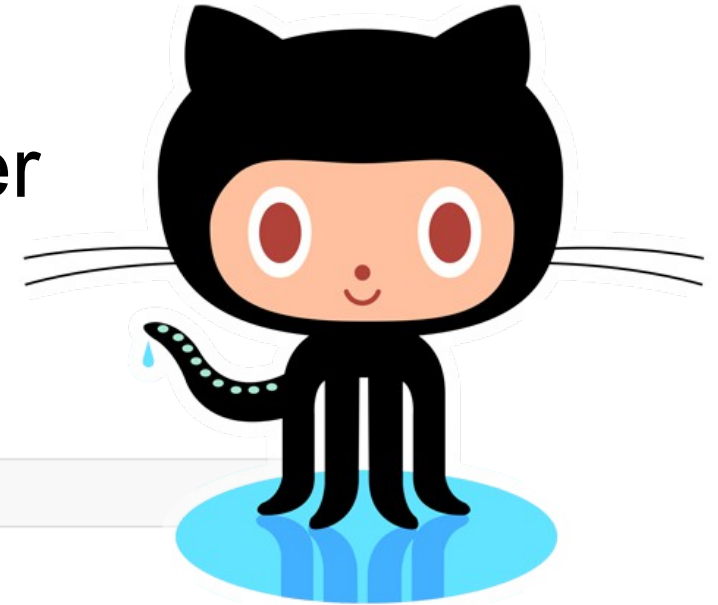
The Abdus Salam
**International Centre
for Theoretical Physics**
50th Anniversary 1964–2014



Repozer unveiled

Get and Install – STEP 1

- <https://github.com/Ictp/Repozer>



STEP 1: Install Repoze staff

- Install Repoze in your Indico instance:

```
$ pip install repoze.catalog
```

- go to `"/src/indico/ext/search/"` and do:

```
git clone https://github.com/Ictp/Repozer.git repozer
```

- Edit `"/src/setup.py"`, add `@566` (right below `"search.invenio = indico.ext.search.invenio"`)

```
search.repozer = indico.ext.search.repozer
```

- Save and from shell in `/src/`:

```
$ python setup.py develop_config
```



Get and Install – STEP 2

STEP 2: Configure and create Catalog

- Start Indico
- Enter Admin web interface, Plugins
- Enable Search plugin, Repozer, set default Sea = Repozer and Save Settings
- Via shell, go to `"/src/indico/ext/search/repozer/manage"` and type:

```
$ python createRepozeCatalog.py
```

this will create the repozecatalog inside your Data.fs (you should see percentage numbers)

- It could take a long time. At the end it will pack your DB.

Now your search engine is up-n-running!



Activate Repozer from Indico's Admin UI

Home | Marie Curie Library | ICTP.tv | Intranet | Webmail | Media Centre

The Abdus Salam
**International Centre
for Theoretical Physics**

Home Create event Server admin My profile Help UTC G. Pieretti

Home » Server Admin

Server Administration

General settings
Users and Groups
IP Domains
Rooms
Layout
Services
Plugins
Homepage
System
Protection

Main **Search**

Active plugins for Search

Enables the search functionality to Indico.

<input checked="" type="checkbox"/> Enabled plugin	<input checked="" type="checkbox"/> Invenio (Invenio Search Engine Agent.)
<input type="checkbox"/> Disabled plugin	<input checked="" type="checkbox"/> Repozer (Repoze Search Engine Agent.)

Reload

Global settings for Search

Default search engine: repozer

Settings for individual plugins

Invenio **Repozer**

No options defined for this plugin

Powered by Indico



Repozer.catalog in action

Home | Marle Curle Library | ICTP.tv | Intranet | Webmail | Media Centre

The Abdus Salam
**International Centre
for Theoretical Physics**

Home Create event Server admin My profile Help UTC G. Pieretti

Search

Search

Search powered by

Hide advanced options

Search in: Title and Description

Category: All

Use wildcards: (will take longer...)

Keywords:

Search for: Events

Start Date:

End Date:

Sort field: Date

Sort order: Descending

Powered by Indico



The result

The screenshot shows the ICTP website's search interface. At the top, there is a navigation bar with links for Home, Marie Curie Library, ICTP.tv, Intranet, Webmail, and Media Centre. Below this is the ICTP logo and name. A secondary navigation bar includes Home, Create event, Server admin, My profile, Help, UTC, and G. Pleretti. The main content area features a search bar with the query 'physics' and a 'Search' button. To the right of the search bar, it says 'Search powered by Repoze'. Below the search bar, the results are displayed as a list of items, each with a title, date, and a brief description. The first result is 'Joint ICTP/SISSA Statistical Physics seminar: "Functional renormalization group and a compactly supported smooth function"' dated 2013-12-18. The second is 'Joint ICTP/SISSA Statistical Physics seminar: "Bringing quantum thermodynamics to the laboratory"' dated 2013-12-10. The third is 'Joint ICTP-VAST-APCTP Regional School and Conference on Theoretical Physics in Topological Phases and Quantum Computation | (smr 2518)' dated 2013-12-09. The fourth is 'Joint ICTP-IAEA International Training Workshop on Accuracy Requirement and Uncertainty in Radiation Therapy | (smr 2502)' dated 2013-12-09. At the bottom of the results, there is a note about the Activity Secretariat: 'The Abdus Salam International Centre for Theoretical Physics Strada Costiera 11 I-34151 Trieste, Italy Telefax:'. The total number of hits is 128.

Home | Marie Curie Library | ICTP.tv | Intranet | Webmail | Media Centre

ICTP The Abdus Salam
International Centre
for Theoretical Physics

Home Create event Server admin My profile Help UTC G. Pleretti

Search

physics Search

Show advanced options Search powered by Repoze

Hits: 128

Joint ICTP/SISSA Statistical Physics seminar: "Functional renormalization group and a compactly supported smooth function"
2013-12-18 00:00:00 (CET)

Functional renormalization group (RG) method is suitable to perform renormalization non-perturbatively. However, the approximated RG equations depend on the choice of the so called regulator function. Thus, optimisation is required in order to increase the predicting power of the functional RG appr...

Joint ICTP/SISSA Statistical Physics seminar: "Bringing quantum thermodynamics to the laboratory"
2013-12-10 00:00:00 (CET)

Research on the out-of-equilibrium dynamics of quantum systems has so far produced important statements on the thermodynamics of small systems undergoing quantum mechanical evolutions. Key examples are provided by the Crooks and Jarzynski relations: taking into account fluctuations in non-equilibri...

Joint ICTP-VAST-APCTP Regional School and Conference on Theoretical Physics in Topological Phases and Quantum Computation | (smr 2518)
2013-12-09 09:00:00 (CET)

Local Secretariat in Vietnam: Ms Duong Thi Man and Dr. Nguyen Thi Kim Thanh Institute of Physics VAST 10 Dao Tan Street, Ba Dinh District, Hanoi, Vietnam Telephone: +84-4-37662107 Telefax: +84-4-37662107 mailto:hoithao@iop.vast.ac.vn Local Secretariat in Trieste: Ms. Marina de Comell...

Joint ICTP-IAEA International Training Workshop on Accuracy Requirement and Uncertainty in Radiation Therapy | (smr 2502)
2013-12-09 08:30:00 (CET)

Activity Secretariat: The Abdus Salam International Centre for Theoretical Physics Strada Costiera 11 I-34151 Trieste, Italy Telefax:



Exporting with Repozer: http_api

- <http://indico.ictp.it/indico/export/conference/search.xml?today=2013/04/10>
 - .xml : export type, can be: .xml/.json/.html/.atom
 - today : if present, return all conferences which startDate <= today AND endDate >= today (running conferences)
- OTHER PARAMETERS:
 - start_date : return conferences where startDate >= start_date
 - end_date : return conferences where endDate <= end_date
 - eg: [..]/search.xml?start_date=2013/01/10
 - eg: [..]/search.xml?start_date=2013/04/10&end_date=2013/04/10
 - NB: start_date and end_date are ALTERNATIVE to today
 - keywords : return confs with specific keyword. Can be a list. OR applied to keyword lists.
 - eg: [..]/search.xml?keywords=Physics and Development, Condensed Matter and Statistical Physics
 - category : return confs with specific category. Can be a list. OR applied to keyword lists.
 - eg: [..]/search.xml?category=RESEARCH IN CONDENSED MATTER PHYSICS, ICTP COLLOQUIUM
 - detail : specify output details level. Default: events
 - Options: events, contributions, subcontributions, sessions
 - eg: <http://indico.ictp.it/indico/export/conference/search.xml?today=2013/04/10&detail=contributions>

http_api: the result

← → ↻ indico.ictp.it/indico/export/conference/search.xml?today=2013/04/10

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<httpapiresult>
  <count>3</count>
  <additionalInfo>
    <collection/>
  </additionalInfo>
  <complete>True</complete>
  <url>/export/conference/search.xml?today=2013/04/10</url>
  <ts>1391199030</ts>
  <results>
    <conference id="a12257" fossil="conferenceMetadata">
      <startDate>2013-04-07T22:00:00+00:00</startDate>
      <endDate>2013-04-12T16:18:00+00:00</endDate>
      <hasAnyProtection>False</hasAnyProtection>
      <keywords>
        <item>smrH323</item>
        <item>expparts70</item>
      </keywords>
      <modificationDate>2014-01-22T14:03:51.380596+00:00</modificationDate>
      <timezone>Europe/Rome</timezone>
      <category>Hosted activities</category>
      <title>
        Theoretical Course "Mouse Genetics; Models for Human Diseases" | (smr H323)
      </title>
      <location>Trieste - Italy</location>
      <type>conference</type>
      <categoryId>21130</categoryId>
      <description/>
      <roomMapURL/>
      <material/>
      <visibility>
        <collection id="">
          <name>Everywhere</name>
        </collection>
      </visibility>
      <address/>
      <room/>
      <roles>
        [{"id":1,"value":"Organizer(s)","editable":false,"child":[{"id":0,"familyName":"ICGEB"}]}]
      </roles>
      <chairs/>
      <url>http://indico.ictp.it/indico/event/a12257</url>
    </conference>
    <conference id="a12167" fossil="conferenceMetadata">
```


http_api with Contributions

```
← → ↻ indico.ictp.it/indico/export/conference/search.xml?today=2013/04/10&detail=contributions
▼<conference id="a12167" fossil="conferenceMetadataWithContribsAndBreaks">
  <startDate>2013-04-08T06:00:00+00:00</startDate>
  <endDate>2013-04-19T11:30:00+00:00</endDate>
  <hasAnyProtection>False</hasAnyProtection>
  <modificationDate>2014-01-27T10:08:54.241020+00:00</modificationDate>
  ▼<keywords>
    <item>smr2446</item>
    <item>expparts65-70</item>
    <item>Miscellaneous</item>
    <item>Policy and Planning</item>
  </keywords>
  <timezone>Europe/Rome</timezone>
  ▼<contributions>
    ▼<contribution id="24" fossil="contributionMetadata">
      <startDate>2013-04-10T09:45:00+00:00</startDate>
      ▼<speakers>
        ▼<contributionparticipation id="0" fossil="contributionParticipationMetadata">
          <affiliation>IAEA</affiliation>
          <fullName>D. LAMBERT</fullName>
          <email/>
        </contributionparticipation>
      </speakers>
      <endDate>2013-04-10T10:30:00+00:00</endDate>
      <description/>
      <title>Exercise: Threat Assessment</title>
      <track/>
      <material/>
      <coauthors/>
      <roomFullname/>
      ▼<url>
        http://indico.ictp.it/indico/event/a12167/session/32/contribution/24
      </url>
      <session>Exercise: Threat Assessment</session>
      <location>Trieste - Italy</location>
      <keywords/>
      <duration>45</duration>
      ▼<primaryauthors>
        ▼<contributionparticipation id="0" fossil="contributionParticipationMetadata">
          <affiliation>IAEA</affiliation>
          <fullName>D. LAMBERT</fullName>
          <email/>
        </contributionparticipation>
      </primaryauthors>
      <type/>
      <room>Leonardo da Vinci Building Main Lecture Hall</room>
    </contribution>
    ▼<contribution id="25" fossil="contributionMetadata">
```



BEWARE!

- Repozer is in **Alpha**:
 - NOT USED IN PRODUCTION YET
- KNOWN BUGS:
 - Index for Conference and Contribution only
(don't care about Lecture, Meeting, SubContributions...)
 - Special characters are not allowed
- SOON (trust me but be patience...):
 - Indexing of Materials (PDF, PPTX, DOCX, XLSX,...)



References

- OUR TEST SITE: <http://indico.ictp.it/indico>
- GET REPOZER: <https://github.com/lctp/Repozer>
... don't forget to fork and contribute!
- REPOZE.CATALOG DOCS: <http://docs.repoze.org/catalog/index.html>
- HOW TO QUERY: <http://docs.repoze.org/catalog/usage.html#searching>
- USAGE EXAMPLE:
 - <http://mainlydata.kubadev.com/python/repoze-catalog-and-zodb-beginners-example-part-1/>
 - <http://mainlydata.kubadev.com/python/repoze-catalog-and-zodb-beginners-example-part-2/>



The Abdus Salam
**International Centre
for Theoretical Physics**
50th Anniversary 1964–2014



Thank you!



Giorgio Pieretti
pieretti@ictp.it